The Exploration of Classification using Neural Networks: Case Study of Souvenir Designs

Komsan Suriya[1] and Sumate Prukrudee[2]

November 2006

Working paper on Artificial Neural Networks
conducted in Chiang Mai, Thailand

[1] Faculty of Economics, Chiang Mai University
and
[2] Social Research Institute, Chiang Mai University
Chiang Mai, Thailand

1. Introduction

After Komsan(2006a) did an experiment on Artificial Neural Networks (ANNs) for classification of souvenir designs, there were rooms for further attempts on experiments for the improvement of the performance of the model. In that paper, he introduced five indicators for the measurement of accuracy; the Elite rate, the Opportunity Lost Rate, the Overall Rate, the Difference between Elite and Opportunity Lost Rate; and the Ratio of not predict anything as successful. In this study, we followed the same measurements but expanded to more variety of structures of ANNs. The experiments were divided into two eras. The first era was before we got suggestions from Assoc. Prof. Krisorn Jittorntrum, our teacher. The second era was after having the suggestions about how we should use the validation set.

All the experiments were done on trial and error basis. We did as many models as we could without a complete plan for the whole paper prior to take-off. It took several months to have enough results to be satisfied. Then we collected the results and put them together into this paper. We hope that they would shed light for us and for other interested people in this field for further improvement of the ANNs.

2. Objectives

2.1 To classify the successful souvenir designs using several structures of ANNs
2.2 To explore the possibility of the improvement of the performance of the ANNs for the classification of souvenir designs

3. Methodologies

First of all, data collection and description of input variables will be briefed in this section. After that, the details of experiments will be provided. The experiments were divided into two era, before and after having the suggestion to avoid over-fitting problem from Dr. Krisorn Jittorntrum.

3.1 Data collection

Catalogues of souvenir designs were launched to international tourists randomly in the domestic departure lounge at Chiang Mai International Airport (CNX) in 2006. There are 285 pictures of souvenir designs in a catalogue. A tourist could choose only two favorite designs. The purpose of choosing is choosing for himself of herself not for other persons. Total frequencies of choosing were 2,552 times from 1,276 tourists.

Successful deigns were those designs that were chosen more than 10 times. The rest was classified as the unsuccessful designs.

3.2 Description of input variables

Eighteen explanatory variables listed in table 1 were included in both Neural Networks and Logit models. The descriptions of all variables are shown in table 1.

Table 1 Description of explanatory variables

	Code	Description
1	HANDMADE	Made by hand
2	WOOD	Made of wood
3	CLOTH	Mode of clothe
4	SILVER	Made of silver
5	NATRMATR	Made of natural materials
6	NATURE	Inspired by nature
7	LANNA	Lanna (Northern Thai) style
8	THAI	Thai style
9	ELEPH	Contains something related to elephant
10	SMALL	Small size
11	LIGHT	Light weight
12	TOLERENC	Not fragile
13	ROUND	Round shape
14	FOLD	foldable
15	PACK	Sold more than 1 piece at a time
16	USE	Function, not just decoration
17	BODY	Usable with human body
18	COLLECT	Collectible as collections

Note: All variables are binary choices with -1 and 1

Although the explanatory variables were binary with -1 and 1, the dependent variable was binary with 0 and 1. The value of one indicated that the design was successful and vice versa.

3.3 Construction of ANNs models

Seven models of ANNs were constructed as listed in Table 2.

Table 2 Structures of all ANNs in this study

Model Name*	**Transformation function**	**Parts**	**Balance Training**	**Training: Validation: Testing**
The first era before having the suggestion about the validation set				
First Model99	tansig,tansig	2	Unbalanced	235:00:50
First Model101	tansig,tansig	2	Unbalanced	235:00:50
Balance55	tansig,logsig	2	Balancd 55:55	110:50:50
The second era after having the suggestion about the validation set				
3parts	tansig,tansig	3	Unbalanced	185:50:50
Stage4	tansig,logsig	3	Unbalanced	185:50:50
Bal46	tansig,logsig	3	Balanced46:46	92:50:50
SelBal46**	tansig,logsig	3	Balanced46:46	92:50:50

Note:

* All models contained only one hidden layer.

** Selctive Learning X Data={-1,1}

The first three models were constructed before having the suggestion about the validation set. They were called the models from the first era. The rest contained validation set after having the suggestions. They were called the models from the second era.

The procedure for finding optimal neurons and epochs were mentioned in next sections.

3.4 The procedure for finding the optimality of the models from the first era

In the first era, each model was simulated using numbers of neurons in the hidden layer ranged from 100 to 1,000 neurons with an additional 100 neurons each time (100, 200,300, …, 1,000 neurons). Each set was simulated 10 times.

The simulation was using the training set as the input. The results obtained from the simulation were compared with the true value in the training set. This was called "in-model" simulation by the authors.

On each time of simulation, five indicators were calculated. They were the Elite rate, the Opportunity Lost Rate, the Overall Rate, the Difference between Elite and Opportunity Lost Rate, and the ratio of predicting anything as successful. The related formulas and commands in Matlab were listed bellowed.

(A) Elite Rate = Percentage of prediction "yes" and the true value is "yes"

(B) Opportunity Lost Rate = Percentage of prediction "no" but the true value is "yes"

(C) Overall Rate = Percentage of "True" prediction out of all predictions

(D) Difference between Elite and Opportunity Lost Rate = Elite Rate – Opp.Lost Rate

(E) Ratio of predicting anything as successful = Frequency of not predict anything as successful divided by all predictions

From ten trails of changing weight initializations, calculate the average of each indicator. Then, calculate the final score of each model using those average values by this following formula.

$$score = \frac{(4*A)+(3*(100-E))+(2*(100-B))+(0.5*C)+(0.5*(100+D))}{4+3+2+0.5+0.5}$$

Select the model with the highest score to be the optimal model for each structure of ANNs.

With the optimal model, process 100 trials of weight initializations. For this time, input values were from the testing set. The results from each simulation were compared with true values in th testing set as well. This was called "out-model" testing by the authors. Calculate five indicators for each trial and then calculate the average value from 100 trials. Finally, calculate the score using the same formula stated above.

3.5 The procedure for finding the optimality of the models from the second era

In the second era, three sets of data were divided; training set, validation set, and testing set. "In-model" testing was not produced in this era anymore since it was better to avoid the over-fitting problem by using the validation set instead. Thus the validation was done by firstly training the model with the training set and simulating using inputs from the validation set. Comparison between the results from each simulation with the true values from the validation set produced an indicator, Mean Squared Error (MSE).

MSE could be calculated by this formula.

$$MSE = \frac{1}{N}\sum_{i=1}^{N} (t_i - y_i)^2$$

Where MSE = Mean Squared Error
N = Number of observations in the validation set
t = True value
y = simulated value
i = index for each observation (i=1,2,…,N)

Another change in the second era was that the number of training round (epoch) was optimized simultaneously with the number of neuron in the hidden layer. Fixing a number of neuron in the hidden layer, 100 rounds (epochs) was trained in the first trial. Each trial produced an MSE.

After that, Train more 100 epochs and calculate another MSE. Repeat this procedure until it reached 1,000 epochs. After all, 10 MSEs were produced from 10 trials.

Repeat these procedures 10 times. After all, 10 MSEs would be produced for a fixed number of neuron and a fixed number of epochs. Calculate the average MSE. It should be noted that within a fixed number of neuron, 10 average MSEs were produced, each from 10 related MSEs. It could be noticed that totally an experiment produced 100 MSEs within a fixed number of neuron.

Varying the number of neuron in the hidden layer, 10 more average MSEs would be produced. The range of variation of the number of neuron in the hidden layer was arbitrary to researchers. In this study, they were usually varied from 100 neurons to 300 neurons with an increment of 10 neurons each time. Some experiments may be differed since researchers may be interested in searching for something along the roadside which laid outside the frame.

After having a complete set of average MSEs from all trials, select the number of neuron as well as the number of epochs which yielded the minimum average MSE to be the best model.

Bring the best model to the final competition. The "out-model" which input variables and true values came from the testing set would be done in this step. Calculate three basic indicators; Elite rate, Opportunity lost rate, and Overall rate. Calculate three more indicators; Ratio of not predicting anything as successful (Ratio Nan), Ratio of Elite rate equals to zero, and Ratio of Opportunity lost rate equals to one hundred.

Ratio NaN = Frequency of not predicting anything as successful divided by all predictions and multiplied by 100

Ratio Elite rate equals to zero =
Frequency of predictions that yielded Elite rate equals to zero divided by all predictions and multiplied by 100

Ratio Opportunity lost rate equals to one hundred =
Frequency of predictions that yielded Opportunity lost rate equals to one hundred divided by all predictions and multiplied by 100

In the final step, calculate the final score by using this formula.

Total score = Positive score – Negative score

Positive score = 0.75*Elite rate + 0.25*Overall Rate

Negative score = 0.25*Opportunity lost rate + 0.25*Ratio NaN
+ 0.25*Ratio Elite rate equals to zero
+ 0.25*Ratio of Opportunity lost rate equals to one hundred

It should be noted that why the authors did not use MAPE (Mean Absolute Percentage Error) which was used by Komsan(2006b) in this study. MAPE was good for measurement of errors in time-series analysis. In the classification problem, the true values were not continuous like in time-series data. They were binary. Thus the measurement of MAPE which concerned on how much were the errors differed from the true value was somehow meaningless in this context. The rates which indicated how many times that the model predicted correctly among all predictions would be more meaningful in the classification work.

4. Findings

The results of all seven models will be shown accordingly to the period of their constructions. In each model, the selection of the best model will be presented first. After that, the performance of the best model in classification will be shown.

4.1 Model selections

4.1.1 Model 1 and 2: FirstModel99 and FirstModel101

According to Komsan(2006a), the best model contained 100 neurons in the hidden layer. In this study, the authors would like to bring the model to the final competition to see how goodness of its performance. However, in the final competition, the model which was called "FirstModel100" could not be simulated 100 rounds of weight initializations (just 16 rounds and hanged). Therefore, its neighbors came instead. The FirstModel99 which contained 100 neurons and the FirstModel101 which contained 101 neurons may be good substitutions for the former best model.

4.1.2 Model 3: Balance55

Table 3 Indicators and score of Model 3 (Balance55) in the model selection process

Neurons In hidden layer	The average values of indicators from 10 trials of weight initializations					Score
	Elite rate	Opp. Lost rate	Overall rate	Diff. b/w Elite-Opp.	Ratio NaN	
100	19.0627	51.1112	53.4	-32.0485	0	53.47
200	22.0017	51.1111	58.8	-29.1094	0	**55.06**
300	19.8037	47.7777	53.2	-27.974	0	54.63
400	19.4947	50	53.8	-30.5053	0	53.96
500	19.4033	52.2223	54.8	-32.819	0	53.42
600	19.5609	49.9999	54	-30.439	0	54.00
700	18.7483	48.8889	51.6	-30.1406	0	53.79
800	19.2291	52.2224	54.4	-32.9933	0	53.32
900	19.4452	51.1111	54.4	-31.6659	0	53.69
1000	18.6251	54.4448	54.2	-35.8197	0	52.48

Source: Calculation using Matlab and Excel

4.1.3 Model 4: 3 Parts

Table 4 MSEs obtained from Model 4 (3 Parts) in the model selection process

Epochs	Number of Neurons in hidden layer					
	1	10	20	30	40	50
100	0.5664	0.5578	0.4966	0.5127	0.5693	0.5257
200	0.5595	0.5166	0.4747	0.5164	0.5503	0.4849
300	0.5561	0.5138	0.5000	0.4892	0.4372	0.4841
400	0.5557	0.5288	0.4645	0.4827	0.4313	0.4793
500	0.5557	0.5440	0.4738	0.4996	0.4630	0.4579
600	0.5557	0.5483	0.4738	0.4868	0.4358	0.4495
700	0.5557	0.5364	0.4795	0.4872	0.4288	0.4519
800	0.5557	0.5291	0.4795	0.4953	0.4336	0.4519
900	0.5557	0.5320	0.4795	0.4952	0.4537	0.4519
1000	0.5557	0.5357	0.4795	0.5011	0.4563	0.4519

Source: Calculation using Matlab

Table 5 MSEs obtained from Model 4 (3 Parts) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	60	70	80	90	100	110
100	0.4941	0.5513	0.6262	0.5886	0.4762	0.5303
200	0.5155	0.5101	0.5085	0.5213	0.4201	0.4879
300	0.4904	0.4890	0.4991	0.5327	0.4334	0.4831
400	0.4857	0.5022	0.5100	0.5615	0.4284	0.4655
500	0.4489	0.4944	0.5082	0.5455	0.4621	0.4608
600	0.4726	0.4997	0.5173	0.5291	0.4579	0.4655
700	0.5031	0.4996	0.5199	0.5305	0.4427	0.4753
800	0.4962	0.5015	0.5198	0.5379	0.4391	0.4722
900	0.4961	0.5015	0.5184	0.5412	0.4391	0.4740
1000	0.4967	0.5015	0.5182	0.5377	0.4384	0.4748

Source: Calculation using Matlab

Table 6 MSEs obtained from Model 4 (3 Parts) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	120	130	**140**	150	160	170
100	0.5116	0.5899	0.4882	0.4362	0.5156	0.5422
200	0.4840	0.5011	0.4771	0.4411	0.4734	0.5179
300	0.5097	0.4861	0.4632	0.4440	0.4928	0.5164
400	0.4826	0.4566	0.4408	0.4470	0.4857	0.5290
500	0.4826	0.4811	0.4325	0.4438	0.5067	0.5158
600	0.4898	0.4684	0.4193	0.4437	0.5029	0.5192
700	0.4857	0.4712	0.4297	0.4437	0.5337	0.5205
800	0.4894	0.4807	**0.4135**	0.4437	0.5336	0.5182
900	0.4905	0.4724	0.4466	0.4437	0.5188	0.5184
1000	0.4891	0.4702	0.4273	0.4437	0.5216	0.5184

Source: Calculation using Matlab

Table 7 MSEs obtained from Model 4 (3 Parts) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	180	190	200	210	220	230
100	0.5376	0.5497	0.6044	0.5675	0.4977	0.5900
200	0.5111	0.5430	0.5406	0.4789	0.4992	0.5395
300	0.5163	0.5028	0.4872	0.4658	0.4986	0.4985
400	0.5046	0.5074	0.4959	0.4788	0.4793	0.5039
500	0.5120	0.5138	0.4987	0.4779	0.4710	0.5055
600	0.5126	0.5078	0.5272	0.4773	0.4714	0.4996
700	0.5126	0.5080	0.5227	0.4773	0.4763	0.5052
800	0.5128	0.5055	0.5217	0.4773	0.4735	0.5051
900	0.5068	0.5055	0.5217	0.4773	0.4676	0.5041
1000	0.5065	0.5055	0.5217	0.4773	0.4716	0.5053

Source: Calculation using Matlab

Table 8 MSEs obtained from Model 4 (3 Parts) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer						
	240	250	260	270	280	290	300
100	0.5065	0.4866	0.6333	0.5387	0.5480	0.5804	0.5935
200	0.4635	0.4378	0.5201	0.5057	0.4714	0.4725	0.4632
300	0.4713	0.4343	0.5136	0.4890	0.4502	0.4536	0.4752
400	0.4514	0.4326	0.5080	0.4899	0.4533	0.4589	0.4730
500	0.4564	0.4423	0.5082	0.4950	0.4541	0.4688	0.4708
600	0.4446	0.4296	0.5082	0.4900	0.4528	0.4607	0.4844
700	0.4758	0.4292	0.5082	0.4974	0.4516	0.4938	0.4701
800	0.4858	0.4254	0.5082	0.4932	0.4516	0.4686	0.4804
900	0.4819	0.4266	0.5082	0.4920	0.4516	0.4877	0.4781
1000	0.4793	0.4266	0.5082	0.4920	0.4516	0.4730	0.4781

Source: Calculation using Matlab

4.1.4 Model 5: Stage4

Table 9 MSEs obtained from Model 5 (Stage4) in the model selection process

Epochs	Number of Neurons in hidden layer					
	1	10	20	30	40	50
100	0.1363	0.1108	0.1157	0.1131	0.1232	0.1035
200	0.1302	0.1204	0.1058	0.1223	0.1064	0.1036
300	0.1296	0.1275	0.1067	0.1135	0.1095	0.1086
400	0.1295	0.1350	0.1079	0.1134	0.1042	0.1044
500	0.1295	0.1359	0.1069	0.1220	0.1050	0.1053
600	0.1327	0.1413	0.1063	0.1269	0.1048	0.1062
700	0.1365	0.1422	0.1062	0.1256	0.1048	0.1062
800	0.1366	0.1421	0.1064	0.1255	0.1048	0.1062
900	0.1366	0.1416	0.1123	0.1255	0.1048	0.1062
1000	0.1366	0.1411	0.1113	0.1255	0.1048	0.1062

Source: Calculation using Matlab

Table 10 MSEs obtained from Model 5 (Stage4) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	60	70	80	90	100	110
100	0.1154	0.1222	0.1308	0.1305	0.1233	0.1462
200	0.1052	0.1161	0.1126	0.1129	0.1259	0.1215
300	0.1062	0.1232	0.1070	0.1057	0.1210	0.1165
400	0.1042	0.1180	0.1160	0.1055	0.1185	0.1174
500	0.1066	0.1168	0.1101	0.1018	0.1227	0.1104
600	0.1099	0.1164	0.1115	0.0999	0.1264	0.1155
700	0.1041	0.1164	0.1087	0.1008	0.1256	0.1158
800	0.1074	0.1164	0.1079	0.1000	0.1266	0.1162
900	0.1078	0.1164	0.1079	0.0977	0.1235	0.1161
1000	0.1073	0.1164	0.1079	0.0991	0.1244	0.1169

Source: Calculation using Matlab

Table 11 MSEs obtained from Model 5 (Stage4) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	120	130	140	150	160	170
100	0.1258	0.1415	0.1340	0.1167	0.1316	0.1296
200	0.1125	0.1328	0.1114	0.1130	0.1232	0.1297
300	0.1148	0.1225	0.1177	0.1142	0.1203	0.1299
400	0.1188	0.1154	0.1204	0.1147	0.1191	0.1235
500	0.1165	0.1193	0.1096	0.1134	0.1208	0.1259
600	0.1192	0.1166	0.1103	0.1138	0.1186	0.1263
700	0.1116	0.1091	0.1159	0.1137	0.1210	0.1257
800	0.1100	0.1121	0.1117	0.1137	0.1220	0.1258
900	0.1101	0.1090	0.1116	0.1137	0.1226	0.1258
1000	0.1093	0.1066	0.1121	0.1137	0.1208	0.1258

Source: Calculation using Matlab

Table 12 MSEs obtained from Model 5 (Stage4) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	180	190	200	210	220	230
100	0.1350	0.1199	0.1345	0.1496	0.1125	0.1337
200	0.1303	0.1156	0.1238	0.1305	0.1113	0.1240
300	0.1221	0.1159	0.1227	0.1140	0.1128	0.1241
400	0.1123	0.1161	0.1218	0.1217	0.1164	0.1225
500	0.1143	0.1162	0.1165	0.1116	0.1133	0.1234
600	0.1159	0.1162	0.1186	0.1185	0.1123	0.1263
700	0.1160	0.1162	0.1245	0.1204	0.1126	0.1247
800	0.1150	0.1162	0.1203	0.1234	0.1126	0.1253
900	0.1152	0.1162	0.1263	0.1250	0.1126	0.1251
1000	0.1162	0.1162	0.1255	0.1267	0.1126	0.1208

Source: Calculation using Matlab

Table 13 MSEs obtained from Model 5 (Stage4) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer						
	240	250	**260**	270	280	290	300
100	0.1123	0.1170	0.1095	0.1331	0.1210	0.1373	0.1219
200	0.1081	0.1169	0.0961	0.1202	0.1174	0.1406	0.1232
300	0.1019	0.1170	0.0977	0.1082	0.1167	0.1256	0.1225
400	0.1059	0.1093	**0.0946**	0.1118	0.1177	0.1097	0.1200
500	0.0997	0.1084	0.0971	0.1091	0.1166	0.1084	0.1261
600	0.0997	0.1114	0.0961	0.1085	0.1154	0.1080	0.1230
700	0.1084	0.1076	0.0964	0.1085	0.1162	0.1086	0.1224
800	0.1097	0.1094	0.0961	0.1085	0.1152	0.1086	0.1199
900	0.1095	0.1089	0.0963	0.1085	0.1150	0.1074	0.1233
1000	0.1106	0.1131	0.0964	0.1085	0.1150	0.1080	0.1238

Source: Calculation using Matlab

4.1.5 Model 6: Bal46

Table 14 MSEs obtained from Model 6 (Bal46) in the model selection process

Epochs	Number of Neurons in hidden layer					
	1	10	20	30	40	50
100	0.2609	0.2978	0.2800	0.2828	0.2765	0.2818
200	0.2637	0.3064	0.2887	0.2899	0.2859	0.2912
300	0.2641	0.3105	0.2923	0.2907	0.2885	0.2947
400	0.2643	0.3125	0.2923	0.2907	0.2885	0.2947
500	0.2644	0.3126	0.2923	0.2907	0.2885	0.2947
600	0.2644	0.3126	0.2923	0.2907	0.2885	0.2947
700	0.2644	0.3126	0.2923	0.2907	0.2885	0.2947
800	0.2644	0.3126	0.2923	0.2907	0.2885	0.2947
900	0.2644	0.3126	0.2923	0.2907	0.2885	0.2947
1000	0.2644	0.3126	0.2923	0.2907	0.2885	0.2947

Source: Calculation using Matlab

Table 15 MSEs obtained from Model 6 (Bal46) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	60	70	80	90	100	110
100	0.2849	0.2862	0.2748	0.3111	0.2874	0.2740
200	0.2964	0.2976	0.2841	0.2971	0.2935	0.2844
300	0.2986	0.3008	0.2875	0.2895	0.2961	0.2888
400	0.2986	0.3008	0.2875	0.2920	0.2960	0.2894
500	0.2986	0.3008	0.2875	0.2923	0.2960	0.2894
600	0.2986	0.3008	0.2875	0.2923	0.2960	0.2894
700	0.2986	0.3008	0.2875	0.2923	0.2960	0.2894
800	0.2986	0.3008	0.2875	0.2923	0.2960	0.2894
900	0.2986	0.3008	0.2875	0.2923	0.2960	0.2894
1000	0.2986	0.3008	0.2875	0.2923	0.2960	0.2894

Source: Calculation using Matlab

Table 16 MSEs obtained from Model 6 (Bal46) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	120	130	140	150	160	170
100	0.2797	0.2826	0.2814	0.2927	0.2850	0.2731
200	0.2898	0.2879	0.2825	0.2945	0.2945	0.2813
300	0.2918	0.2894	0.2864	0.2923	0.2958	0.2823
400	0.2918	0.2894	0.2875	0.2934	0.2958	0.2823
500	0.2918	0.2894	0.2875	0.2934	0.2958	0.2823
600	0.2918	0.2894	0.2875	0.2934	0.2958	0.2823
700	0.2918	0.2894	0.2875	0.2934	0.2958	0.2823
800	0.2918	0.2894	0.2875	0.2934	0.2958	0.2823
900	0.2918	0.2894	0.2875	0.2934	0.2958	0.2823
1000	0.2918	0.2894	0.2875	0.2934	0.2958	0.2823

Source: Calculation using Matlab

Table 17 MSEs obtained from Model 6 (Bal46) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	180	190	200	210	220	230
100	0.2823	0.2694	0.2772	0.2728	0.2936	0.2922
200	0.2895	0.2807	0.2891	0.2830	0.3050	0.3034
300	0.2894	0.2831	0.2917	0.2872	0.3062	0.2971
400	0.2896	0.2830	0.2917	0.2878	0.3067	0.2922
500	0.2896	0.2834	0.2917	0.2878	0.3067	0.2980
600	0.2896	0.2834	0.2917	0.2878	0.3067	0.2961
700	0.2896	0.2834	0.2917	0.2878	0.3067	0.2957
800	0.2896	0.2834	0.2917	0.2878	0.3067	0.2957
900	0.2896	0.2834	0.2917	0.2878	0.3067	0.2957
1000	0.2896	0.2834	0.2917	0.2878	0.3067	0.2957

Source: Calculation using Matlab

Table 18 MSEs obtained from Model 6 (Bal46) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	240	250	260	270	280	290
100	0.2813	0.2743	0.2947	0.2729	0.2802	0.2830
200	0.2914	0.2849	0.2930	0.2809	0.2846	0.2927
300	0.2977	0.2878	0.3016	0.2855	0.2865	0.2939
400	0.2977	0.2878	0.3016	0.2855	0.2865	0.2939
500	0.2977	0.2878	0.3016	0.2855	0.2865	0.2939
600	0.2977	0.2878	0.3016	0.2855	0.2865	0.2939
700	0.2977	0.2878	0.3016	0.2855	0.2865	0.2939
800	0.2977	0.2878	0.3016	0.2855	0.2865	0.2939
900	0.2977	0.2878	0.3016	0.2855	0.2865	0.2939
1000	0.2977	0.2878	0.3016	0.2855	0.2865	0.2939

Source: Calculation using Matlab

Table 19 MSEs obtained from Model 6 (Bal46) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	300	310	320	330	340	350
100	**0.2638**	0.2885	0.2793	0.2701	0.2757	0.2777
200	0.2798	0.3016	0.2884	0.2797	0.2850	0.2873
300	0.2814	0.3044	0.2908	0.2820	0.2881	0.2893
400	0.2815	0.3044	0.2908	0.2826	0.2881	0.2893
500	0.2815	0.3044	0.2908	0.2826	0.2881	0.2893
600	0.2815	0.3044	0.2908	0.2826	0.2881	0.2893
700	0.2815	0.3044	0.2908	0.2826	0.2881	0.2893
800	0.2815	0.3044	0.2908	0.2826	0.2881	0.2893
900	0.2815	0.3044	0.2908	0.2826	0.2881	0.2893
1000	0.2815	0.3044	0.2908	0.2826	0.2881	0.2893

Source: Calculation using Matlab

Table 20 MSEs obtained from Model 6 (Bal46) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer				
	360	370	380	390	400
100	0.2840	0.2825	0.2938	0.2848	0.2739
200	0.2929	0.2949	0.2967	0.2845	0.2849
300	0.2944	0.2964	0.2985	0.2886	0.2859
400	0.2944	0.2964	0.2985	0.2890	0.2859
500	0.2944	0.2964	0.2985	0.2890	0.2859
600	0.2944	0.2964	0.2985	0.2890	0.2859
700	0.2944	0.2964	0.2985	0.2890	0.2859
800	0.2944	0.2964	0.2985	0.2890	0.2859
900	0.2944	0.2964	0.2985	0.2890	0.2859
1000	0.2944	0.2964	0.2985	0.2890	0.2859

Source: Calculation using Matlab

4.1.6 Model 7: SelBal46

Table 21 MSEs obtained from Model 7 (SelBal46) in the model selection process

Epochs	Number of Neurons in hidden layer					
	1	**10**	20	30	40	50
100	0.2310	0.2299	0.2342	0.2440	0.2342	0.2381
200	0.2307	**0.2296**	0.2306	0.2438	0.2352	0.2416
300	0.2307	0.2296	0.2306	0.2438	0.2352	0.2416
400	0.2307	0.2296	0.2306	0.2438	0.2352	0.2416
500	0.2307	0.2296	0.2306	0.2438	0.2352	0.2416
600	0.2307	0.2296	0.2306	0.2438	0.2352	0.2416
700	0.2307	0.2296	0.2306	0.2438	0.2352	0.2416
800	0.2307	0.2296	0.2306	0.2438	0.2352	0.2416
900	0.2307	0.2296	0.2306	0.2438	0.2352	0.2416
1000	0.2307	0.2296	0.2306	0.2438	0.2352	0.2416

Source: Calculation using Matlab

Table 22 MSEs obtained from Model 7 (SelBal46) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	60	70	80	90	100	110
100	0.2349	0.2386	0.2859	0.2451	0.2454	0.2949
200	0.2378	0.2395	0.2926	0.2404	0.2501	0.2920
300	0.2381	0.2395	0.2892	0.2385	0.2501	0.2920
400	0.2381	0.2395	0.2892	0.2385	0.2501	0.2920
500	0.2381	0.2395	0.2892	0.2385	0.2501	0.2920
600	0.2381	0.2395	0.2892	0.2385	0.2501	0.2920
700	0.2381	0.2395	0.2892	0.2385	0.2501	0.2920
800	0.2381	0.2395	0.2892	0.2385	0.2501	0.2920
900	0.2381	0.2395	0.2892	0.2385	0.2501	0.2920
1000	0.2381	0.2395	0.2892	0.2385	0.2501	0.2920

Source: Calculation using Matlab

Table 23 MSEs obtained from Model 7 (SelBal46) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	120	130	140	150	160	170
100	0.2883	0.2890	0.2389	0.2793	0.2384	0.2456
200	0.2481	0.2891	0.2444	0.2775	0.2362	0.2477
300	0.2460	0.2891	0.2378	0.2767	0.2375	0.2466
400	0.2460	0.2891	0.2378	0.2378	0.2375	0.2466
500	0.2460	0.2891	0.2378	0.2378	0.2375	0.2466
600	0.2460	0.2891	0.2378	0.2378	0.2375	0.2466
700	0.2460	0.2891	0.2378	0.2378	0.2375	0.2466
800	0.2460	0.2891	0.2378	0.2378	0.2375	0.2466
900	0.2460	0.2891	0.2378	0.2378	0.2375	0.2466
1000	0.2460	0.2891	0.2378	0.2378	0.2375	0.2466

Source: Calculation using Matlab

Table 24 MSEs obtained from Model 7 (SelBal46) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer					
	180	190	200	210	220	230
100	0.2439	0.2447	0.2480	0.2366	0.3227	0.2309
200	0.2454	0.2430	0.2482	0.2358	0.3228	0.2303
300	0.2440	0.2387	0.2477	0.2358	0.3228	0.2300
400	0.2440	0.2390	0.2477	0.2358	0.3228	0.2300
500	0.2440	0.2390	0.2477	0.2358	0.3228	0.2300
600	0.2440	0.2390	0.2477	0.2358	0.3228	0.2300
700	0.2440	0.2390	0.2477	0.2358	0.3228	0.2300
800	0.2440	0.2390	0.2477	0.2358	0.3228	0.2300
900	0.2440	0.2390	0.2477	0.2358	0.3228	0.2300
1000	0.2440	0.2390	0.2477	0.2358	0.3228	0.2300

Source: Calculation using Matlab

Table 25 MSEs obtained from Model 7 (SelBal46) in the model selection process (cont.)

Epochs	Number of Neurons in hidden layer						
	240	250	260	270	280	290	300
100	0.2395	0.2336	0.3315	0.2470	0.2305	0.2413	0.2955
200	0.2420	0.2369	0.3336	0.2406	0.2332	0.2394	0.2954
300	0.2405	0.2374	0.3336	0.2407	0.2305	0.2394	0.2954
400	0.2405	0.2364	0.3336	0.2407	0.2305	0.2394	0.2954
500	0.2405	0.2364	0.3336	0.2407	0.2305	0.2394	0.2954
600	0.2405	0.2364	0.3336	0.2407	0.2305	0.2394	0.2954
700	0.2405	0.2364	0.3336	0.2407	0.2305	0.2394	0.2954
800	0.2405	0.2364	0.3336	0.2407	0.2305	0.2394	0.2954
900	0.2405	0.2364	0.3336	0.2407	0.2305	0.2394	0.2954
1000	0.2405	0.2364	0.3336	0.2407	0.2305	0.2394	0.2954

Source: Calculation using Matlab

4.1.7 Summary of model selections

Table 26 The summary of the best model which brought to the final comptition

Model Name*	Neurons	Epochs	Transformation function	Parts	Balance Training	Training: Validation: Testing
The first era before having the suggestion about the validation set						
First Model99	99	1000	tansig,tansig	2	Unbalanced	235:00:50
First Model101	101	1000	tansig,tansig	2	Unbalanced	235:00:50
Balance55	200	1000	tansig,logsig	2	Balancd 55:55	110:50:50
The second era after having the suggestion about the validation set						
3parts	140	800	tansig,tansig	3	Unbalanced	185:50:50
Stage4	260	400	tansig,logsig	3	Unbalanced	185:50:50
Bal46	300	100	tansig,logsig	3	Balanced46:46	92:50:50
SelBal46**	10	200	tansig,logsig	3	Balanced46:46	92:50:50

Note:

* All models contained only one hidden layer.

** Selctive Learning X Data={-1,1}

4.2 Final competition

4.2.1 Model 1: FirstModel99

Table 27 Indicators from prediction using Model 1 (FirstModel99) in the final competition

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
1	NaN	100	82	**17%**	**38%**	**55%**
2	NaN	100	82			
3	40	77.778	80			
4	0	100	74			
5	25	88.889	78			
6	0	100	78			
7	50	66.667	82			
8	50	77.778	82			
9	50	77.778	82			
10	0	100	76			
11	50	77.778	82			
12	0	100	78			
13	0	100	74			
14	NaN	100	82			
15	100	88.889	84			
16	40	77.778	80			
17	50	77.778	82			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
18	0	100	80			
19	0	100	78			
20	0	100	78			
21	0	100	78			
22	50	77.778	82			
23	0	100	80			
24	0	100	74			
25	60	66.667	84			
26	50	77.778	82			
27	0	100	78			
28	NaN	100	82			
29	0	100	78			
30	40	77.778	80			
31	33.333	77.778	78			
32	NaN	100	82			
33	60	66.667	84			
34	0	100	80			
35	50	88.889	82			
36	75	66.667	86			
37	50	66.667	82			
38	0	100	78			
39	0	100	78			
40	0	100	78			
41	60	66.667	84			
42	50	77.778	82			
43	NaN	100	82			
44	50	66.667	82			
45	0	100	76			
46	NaN	100	82			
47	50	77.778	82			
48	0	100	76			
49	50	66.667	82			
50	50	77.778	82			
51	33.333	88.889	80			
52	66.667	77.778	84			
53	0	100	76			
54	50	77.778	82			
55	NaN	100	82			
56	50	77.778	82			
57	50	77.778	82			
58	0	100	78			
59	0	100	80			
60	0	100	80			
61	NaN	100	82			
62	NaN	100	82			
63	33.333	77.778	78			
64	0	100	78			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
65	0	100	78			
66	50	66.667	82			
67	NaN	100	82			
68	0	100	78			
69	60	66.667	84			
70	0	100	78			
71	60	66.667	84			
72	50	66.667	82			
73	0	100	78			
74	50	77.778	82			
75	0	100	76			
76	NaN	100	82			
77	50	66.667	82			
78	40	77.778	80			
79	0	100	80			
80	0	100	78			
81	0	100	80			
82	NaN	100	82			
83	0	100	80			
84	NaN	100	82			
85	66.667	77.778	84			
86	50	77.778	82			
87	0	100	78			
88	0	100	80			
89	0	100	80			
90	0	100	74			
91	40	77.778	80			
92	100	77.778	86			
93	NaN	100	82			
94	50	77.778	82			
95	NaN	100	82			
96	50	77.778	82			
97	0	100	78			
98	50	77.778	82			
99	NaN	100	82			
100	40	77.778	80			
Average	**27.99**	**89.00**	**80.40**			

Source: Calculation using Matlab and Excel

4.2.2 Model 2: FirstModel101

Table 28 Indicators from prediction using Model2 (FirstModel101) in the final competition

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
1	1	66.667	77.778	**26%**	**27%**	**53%**
2	2	NaN	100			
3	3	50	77.778			
4	4	50	77.778			
5	5	NaN	100			
6	6	0	100			
7	7	50	77.778			
8	8	0	100			
9	9	NaN	100			
10	10	66.667	77.778			
11	11	0	100			
12	12	50	77.778			
13	13	NaN	100			
14	14	0	100			
15	15	66.667	77.778			
16	16	NaN	100			
17	17	50	77.778			
18	18	0	100			
19	19	0	100			
20	20	0	100			
21	21	50	66.667			
22	22	NaN	100			
23	23	50	77.778			
24	24	0	100			
25	25	50	77.778			
26	26	0	100			
27	27	0	100			
28	28	50	77.778			
29	29	50	88.889			
30	30	0	100			
31	31	60	66.667			
32	32	NaN	100			
33	33	NaN	100			
34	34	0	100			
35	35	0	100			
36	36	0	100			
37	37	NaN	100			
38	38	NaN	100			
39	39	NaN	100			
40	40	16.667	88.889			
41	41	50	77.778			
42	42	0	100			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
43	43	66.667	77.778			
44	44	0	100			
45	45	0	100			
46	46	50	66.667			
47	47	40	77.778			
48	48	0	100			
49	49	0	100			
50	50	50	77.778			
51	51	NaN	100			
52	52	0	100			
53	53	NaN	100			
54	54	66.667	55.556			
55	55	40	77.778			
56	56	0	100			
57	57	60	66.667			
58	58	40	77.778			
59	59	50	66.667			
60	60	0	100			
61	61	50	77.778			
62	62	50	77.778			
63	63	33.333	88.889			
64	64	50	88.889			
65	65	NaN	100			
66	66	NaN	100			
67	67	NaN	100			
68	68	33.333	77.778			
69	69	50	77.778			
70	70	NaN	100			
71	71	60	66.667			
72	72	NaN	100			
73	73	0	100			
74	74	NaN	100			
75	75	50	77.778			
76	76	50	77.778			
77	77	NaN	100			
78	78	0	100			
79	79	NaN	100			
80	80	100	77.778			
81	81	NaN	100			
82	82	0	100			
83	83	60	66.667			
84	84	60	66.667			
85	85	50	77.778			
86	86	60	66.667			
87	87	NaN	100			
88	88	33.333	77.778			
89	89	NaN	100			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
90	90	50	66.667			
91	91	100	77.778			
92	92	50	77.778			
93	93	NaN	100			
94	94	NaN	100			
95	95	50	77.778			
96	96	50	66.667			
97	97	0	100			
98	98	33.333	77.778			
99	99	50	77.778			
100	100	0	100			
Average	**33.29**	**88.56**	**81.10**			

Source: Calculation using Matlab and Excel

4.2.3 Model 3: Balance55

Table 29 Indicators from prediction using Model3 (Balance55) in the final competition

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
1	23.81	44.444	60	**0%**	**0%**	**0%**
2	17.391	55.556	52			
3	18.182	55.556	54			
4	13.043	66.667	48			
5	18.182	55.556	54			
6	19.048	55.556	56			
7	23.81	44.444	60			
8	21.739	44.444	56			
9	30	66.667	74			
10	23.81	44.444	60			
11	19.048	55.556	56			
12	19.048	55.556	56			
13	22.727	44.444	58			
14	21.739	44.444	56			
15	19.048	55.556	56			
16	16	55.556	48			
17	18.182	55.556	54			
18	20	44.444	52			
19	16	55.556	48			
20	17.391	55.556	52			
21	21.739	44.444	56			
22	18.182	55.556	54			
23	19.048	55.556	56			
24	19.048	55.556	56			
25	17.391	55.556	52			
26	20	44.444	52			
27	16.667	55.556	50			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
28	18.182	55.556	54			
29	20	44.444	52			
30	18.182	55.556	54			
31	23.81	44.444	60			
32	19.048	55.556	56			
33	20.833	44.444	54			
34	21.739	44.444	56			
35	20	55.556	58			
36	17.391	55.556	52			
37	20.833	44.444	54			
38	16	55.556	48			
39	20	44.444	52			
40	15.385	55.556	46			
41	21.739	44.444	56			
42	17.391	55.556	52			
43	18.182	55.556	54			
44	20	44.444	52			
45	20.833	44.444	54			
46	19.048	55.556	56			
47	18.182	55.556	54			
48	17.391	55.556	52			
49	14.815	55.556	44			
50	20.833	44.444	54			
51	16	55.556	48			
52	17.391	55.556	52			
53	16.667	55.556	50			
54	17.391	55.556	52			
55	21.739	44.444	56			
56	16	55.556	48			
57	17.391	55.556	52			
58	21.053	55.556	60			
59	20.833	44.444	54			
60	20.833	44.444	54			
61	18.182	55.556	54			
62	25	44.444	62			
63	17.391	55.556	52			
64	19.048	55.556	56			
65	20	44.444	52			
66	21.739	44.444	56			
67	18.182	55.556	54			
68	19.048	55.556	56			
69	17.391	55.556	52			
70	17.391	55.556	52			
71	22.727	44.444	58			
72	15.789	66.667	56			
73	16	55.556	48			
74	19.048	55.556	56			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
75	21.739	44.444	56			
76	20.833	44.444	54			
77	20	44.444	52			
78	21.739	44.444	56			
79	20	44.444	52			
80	19.048	55.556	56			
81	16.667	55.556	50			
82	20.833	44.444	54			
83	17.391	55.556	52			
84	21.739	44.444	56			
85	16.667	55.556	50			
86	21.739	44.444	56			
87	19.048	55.556	56			
88	19.048	55.556	56			
89	19.048	55.556	56			
90	20.833	44.444	54			
91	16	55.556	48			
92	18.182	55.556	54			
93	21.739	44.444	56			
94	16.667	55.556	50			
95	20	55.556	58			
96	22.727	44.444	58			
97	20	44.444	52			
98	20	44.444	52			
99	16.667	55.556	50			
100	18.519	44.444	48			
Average	**19.29**	**51.44**	**53.92**			

Source: Calculation using Matlab and Excel

4.2.4 Model 4: 3 Parts

Table 30 Indicators from prediction using Model4 (3 Parts) in the final competition

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
1	0	100	74	**27%**	**27%**	**54%**
2	0	100	78			
3	NaN	100	82			
4	60	66.667	84			
5	0	100	80			
6	42.857	66.667	80			
7	0	100	80			
8	50	77.778	82			
9	0	100	78			
10	50	77.778	82			
11	40	77.778	80			
12	100	77.778	86			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
13	NaN	100	82			
14	0	100	78			
15	NaN	100	82			
16	12.5	88.889	70			
17	0	100	80			
18	NaN	100	82			
19	0	100	76			
20	50	77.778	82			
21	NaN	100	82			
22	50	77.778	82			
23	0	100	76			
24	33.333	77.778	78			
25	0	100	78			
26	NaN	100	82			
27	66.667	77.778	84			
28	0	100	80			
29	0	100	80			
30	NaN	100	82			
31	0	100	78			
32	0	100	78			
33	0	100	80			
34	50	77.778	82			
35	0	100	76			
36	60	66.667	84			
37	0	100	80			
38	40	77.778	80			
39	42.857	66.667	80			
40	0	100	76			
41	NaN	100	82			
42	0	100	80			
43	40	77.778	80			
44	0	100	76			
45	33.333	77.778	78			
46	33.333	77.778	78			
47	50	77.778	82			
48	50	77.778	82			
49	NaN	100	82			
50	0	100	80			
51	25	88.889	78			
52	0	100	78			
53	40	77.778	80			
54	NaN	100	82			
55	NaN	100	82			
56	33.333	77.778	78			
57	0	100	76			
58	28.571	77.778	76			
59	NaN	100	82			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
60	NaN	100	82			
61	NaN	100	82			
62	40	77.778	80			
63	50	77.778	82			
64	0	100	78			
65	33.333	77.778	78			
66	50	77.778	82			
67	NaN	100	82			
68	40	77.778	80			
69	40	77.778	80			
70	40	77.778	80			
71	NaN	100	82			
72	0	100	76			
73	NaN	100	82			
74	100	77.778	86			
75	50	77.778	82			
76	NaN	100	82			
77	40	77.778	80			
78	50	66.667	82			
79	NaN	100	82			
80	NaN	100	82			
81	50	77.778	82			
82	37.5	66.667	78			
83	NaN	100	82			
84	42.857	66.667	80			
85	NaN	100	82			
86	50	77.778	82			
87	40	77.778	80			
88	60	66.667	84			
89	66.667	77.778	84			
90	NaN	100	82			
91	NaN	100	82			
92	66.667	77.778	84			
93	NaN	100	82			
94	NaN	100	82			
95	50	77.778	82			
96	0	100	80			
97	NaN	100	82			
98	0	100	80			
99	33.333	88.889	80			
100	50	77.778	82			
Average	**29.62**	**89.22**	**80.42**			

Source: Calculation using Matlab and Excel

4.2.5 Model 5: Stage4

Table 31 Indicators from prediction using Model5 (Stage4) in the final competition

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
1	66.667	77.778	84	**16%**	**32%**	**48%**
2	50	77.778	82			
3	40	77.778	80			
4	50	77.778	82			
5	33.333	77.778	78			
6	0	100	80			
7	66.667	77.778	84			
8	50	77.778	82			
9	50	77.778	82			
10	40	77.778	80			
11	0	100	80			
12	0	100	78			
13	50	77.778	82			
14	NaN	100	82			
15	50	77.778	82			
16	40	77.778	80			
17	25	88.889	78			
18	50	77.778	82			
19	NaN	100	82			
20	50	77.778	82			
21	0	100	76			
22	0	100	78			
23	NaN	100	82			
24	NaN	100	82			
25	40	77.778	80			
26	40	77.778	80			
27	42.857	66.667	80			
28	37.5	66.667	78			
29	0	100	78			
30	50	77.778	82			
31	NaN	100	82			
32	33.333	77.778	78			
33	0	100	80			
34	NaN	100	82			
35	40	77.778	80			
36	33.333	77.778	78			
37	60	66.667	84			
38	33.333	77.778	78			
39	0	100	80			
40	0	100	80			
41	33.333	77.778	78			
42	0	100	74			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
43	0	100	80			
44	NaN	100	82			
45	40	77.778	80			
46	0	100	80			
47	0	100	80			
48	50	66.667	82			
49	0	100	80			
50	40	77.778	80			
51	0	100	76			
52	0	100	80			
53	50	77.778	82			
54	0	100	80			
55	0	100	74			
56	40	77.778	80			
57	NaN	100	82			
58	NaN	100	82			
59	0	100	80			
60	0	100	78			
61	0	100	78			
62	50	77.778	82			
63	50	77.778	82			
64	50	77.778	82			
65	50	77.778	82			
66	0	100	76			
67	33.333	77.778	78			
68	80	55.556	88			
69	40	77.778	80			
70	42.857	66.667	80			
71	0	100	80			
72	0	100	76			
73	40	77.778	80			
74	0	100	78			
75	0	100	80			
76	NaN	100	82			
77	NaN	100	82			
78	50	77.778	82			
79	50	77.778	82			
80	NaN	100	82			
81	0	100	76			
82	NaN	100	82			
83	50	77.778	82			
84	NaN	100	82			
85	0	100	80			
86	50	66.667	82			
87	0	100	74			
88	40	77.778	80			
89	50	66.667	82			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
90	NaN	100	82			
91	0	100	76			
92	60	66.667	84			
93	0	100	76			
94	NaN	100	82			
95	50	77.778	82			
96	60	66.667	84			
97	60	66.667	84			
98	0	100	74			
99	66.667	77.778	84			
100	50	77.778	82			
Average	**29.15**	**87.22**	**80.32**			

Source: Calculation using Matlab and Excel

4.2.6 Model 6: Bal46

Table 32 Indicators from prediction using Model5 (Stage4) in the final competition

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
1	28.571	55.556	70	**0%**	**0%**	**0%**
2	20.513	11.111	36			
3	26.667	55.556	68			
4	25	66.667	70			
5	28.571	55.556	70			
6	30.769	55.556	72			
7	28.571	55.556	70			
8	26.667	55.556	68			
9	26.667	55.556	68			
10	23.077	66.667	68			
11	25	55.556	66			
12	40	55.556	78			
13	23.077	66.667	68			
14	33.333	55.556	74			
15	22.222	55.556	62			
16	26.667	55.556	68			
17	25	55.556	66			
18	17.241	44.444	44			
19	23.529	55.556	64			
20	33.333	55.556	74			
21	26.667	55.556	68			
22	36.364	55.556	76			
23	26.667	55.556	68			
24	25	55.556	66			
25	26.667	55.556	68			
26	30.769	55.556	72			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
27	30.769	55.556	72			
28	28.571	55.556	70			
29	36.364	55.556	76			
30	20	44.444	52			
31	33.333	55.556	74			
32	28.571	55.556	70			
33	28.571	55.556	70			
34	33.333	55.556	74			
35	36.364	55.556	76			
36	36.364	55.556	76			
37	33.333	55.556	74			
38	17.647	66.667	60			
39	28.571	55.556	70			
40	28.571	55.556	70			
41	28.571	55.556	70			
42	30.769	55.556	72			
43	26.667	55.556	68			
44	22.222	55.556	62			
45	36.364	55.556	76			
46	25	55.556	66			
47	26.667	55.556	68			
48	33.333	55.556	74			
49	33.333	55.556	74			
50	25	55.556	66			
51	23.529	55.556	64			
52	33.333	55.556	74			
53	26.667	55.556	68			
54	30.769	55.556	72			
55	28.571	55.556	70			
56	25	55.556	66			
57	28.571	55.556	70			
58	23.529	55.556	64			
59	28.571	55.556	70			
60	30.769	55.556	72			
61	28.571	55.556	70			
62	28.571	55.556	70			
63	26.667	55.556	68			
64	36.364	55.556	76			
65	30.769	55.556	72			
66	25	55.556	66			
67	26.667	55.556	68			
68	30.769	55.556	72			
69	26.667	55.556	68			
70	28.571	55.556	70			
71	26.667	55.556	68			
72	36.364	55.556	76			
73	26.667	55.556	68			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
74	36.364	55.556	76			
75	26.667	55.556	68			
76	26.667	55.556	68			
77	23.077	66.667	68			
78	26.667	55.556	68			
79	30.769	55.556	72			
80	23.529	55.556	64			
81	28.571	55.556	70			
82	25	55.556	66			
83	25	55.556	66			
84	33.333	55.556	74			
85	26.667	55.556	68			
86	23.529	55.556	64			
87	36.364	55.556	76			
88	23.81	44.444	60			
89	28.571	55.556	70			
90	30.769	55.556	72			
91	26.667	55.556	68			
92	26.667	55.556	68			
93	30.769	55.556	72			
94	30.769	55.556	72			
95	26.667	55.556	68			
96	28.571	55.556	70			
97	23.529	55.556	64			
98	36.364	55.556	76			
99	30.769	55.556	72			
100	30.769	55.556	72			
Average	**28.39**	**55.33**	**68.96**			

Source: Calculation using Matlab and Excel

4.2.7 Model 7: SelBal46

Table 33 Indicators from prediction using Model5 (Stage4) in the final competition

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
1	50	88.889	82	**0%**	**0%**	**0%**
2	50	88.889	82			
3	33.333	88.889	80			
4	50	88.889	82			
5	50	88.889	82			
6	33.333	88.889	80			
7	50	88.889	82			
8	33.333	88.889	80			
9	50	88.889	82			
10	50	88.889	82			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
11	50	88.889	82			
12	50	88.889	82			
13	33.333	88.889	80			
14	50	88.889	82			
15	33.333	88.889	80			
16	33.333	88.889	80			
17	50	88.889	82			
18	33.333	88.889	80			
19	50	88.889	82			
20	50	88.889	82			
21	20	88.889	76			
22	16.667	88.889	74			
23	33.333	88.889	80			
24	50	88.889	82			
25	20	88.889	76			
26	50	88.889	82			
27	50	88.889	82			
28	50	88.889	82			
29	50	88.889	82			
30	33.333	77.778	78			
31	50	88.889	82			
32	33.333	88.889	80			
33	50	88.889	82			
34	50	88.889	82			
35	33.333	88.889	80			
36	50	88.889	82			
37	20	88.889	76			
38	16.667	88.889	74			
39	50	88.889	82			
40	50	88.889	82			
41	33.333	88.889	80			
42	20	88.889	76			
43	50	88.889	82			
44	33.333	88.889	80			
45	33.333	88.889	80			
46	50	88.889	82			
47	50	88.889	82			
48	50	88.889	82			
49	16.667	88.889	74			
50	50	88.889	82			
51	50	88.889	82			
52	50	88.889	82			
53	50	88.889	82			
54	50	88.889	82			
55	33.333	88.889	80			
56	33.333	88.889	80			
57	50	88.889	82			

Trials of weight initializations	Elite	Opp	Overall	Ratio NaN	Ratio Elite=0	Ratio Opp.=100
58	50	77.778	82			
59	33.333	88.889	80			
60	33.333	88.889	80			
61	16.667	88.889	74			
62	33.333	88.889	80			
63	50	88.889	82			
64	33.333	88.889	80			
65	50	88.889	82			
66	20	88.889	76			
67	33.333	88.889	80			
68	50	88.889	82			
69	16.667	88.889	74			
70	33.333	88.889	80			
71	50	88.889	82			
72	50	88.889	82			
73	33.333	88.889	80			
74	20	88.889	76			
75	50	88.889	82			
76	50	88.889	82			
77	33.333	88.889	80			
78	33.333	88.889	80			
79	50	88.889	82			
80	16.667	88.889	74			
81	50	88.889	82			
82	33.333	88.889	80			
83	50	88.889	82			
84	33.333	88.889	80			
85	50	88.889	82			
86	50	88.889	82			
87	50	88.889	82			
88	16.667	88.889	74			
89	50	88.889	82			
90	33.333	88.889	80			
91	33.333	88.889	80			
92	33.333	88.889	80			
93	50	88.889	82			
94	33.333	88.889	80			
95	50	88.889	82			
96	33.333	88.889	80			
97	33.333	88.889	80			
98	16.667	88.889	74			
99	33.333	88.889	80			
100	50	77.778	82			
Average	**39.87**	**88.56**	**80.30**			

Source: Calculation using Matlab and Excel

4.2.8 Summary of final competition

Table 34 Summary of average values of indicators and ratios

Model	Average value of indicators			Ratio NaN	Ratio Elite=0	Ratio Opp.=100
	Elite rate	Opp. rate	Overall rate			
FirstModel99	27.99	89.00	80.40	17	38	55
FirstModel101	33.29	88.56	81.10	26	27	53
Balance55	19.29	51.44	53.92	0	0	0
3parts	29.62	89.22	80.42	27	27	54
Stage4	29.15	87.22	80.32	16	32	48
Bal46	28.39	55.33	68.96	0	0	0
SelBal46	39.87	88.56	80.30	0	0	0

Source: Calculation using Matlab and Excel

Table 35 Positive score of models

Models	0.75*Elite rate	0.25*Overall rate	Positive score
FirstModel99	20.99	20.10	**41.09**
FirstModel101	24.97	20.28	**45.24**
Balance55	14.47	13.48	**27.95**
3parts	22.21	20.11	**42.32**
Stage4	21.86	20.08	**41.94**
Bal46	21.30	17.24	**38.54**
SelBal46	29.90	20.08	**49.97**

Source: Calculation using Excel

Table 36 Negative score of models

Models	0.25*Opp	025*Na/N	0.25*(Elite=0)	0.25*(Opp=100)	Negative Score
FirstModel99	22.25	4.25	9.5	13.75	**49.75**
FirstModel101	22.14	6.5	6.75	13.25	**48.64**
Balance55	12.86	0	0	0	**12.86**
3parts	22.31	6.75	6.75	13.5	**49.31**
Stage4	21.81	4	8	12	**45.81**
Bal46	13.83	0	0	0	**13.83**
SelBal46	22.14	0	0	0	**22.14**

Source: Calculation using Excel

Table 37 Total score of models

Models	Positive score	Negative score	Total score
FirstModel99	41.09	49.75	-8.66
FirstModel101	45.24	48.64	-3.40
Balance55	27.95	12.86	15.09
3parts	42.32	49.31	-6.99
Stage4	41.94	45.81	-3.87
Bal46	38.54	13.83	24.70
SelBal46	49.97	22.14	27.84

Source: Calculation using Excel

5. Conclusions

In this study, several structures of ANNs were done for the explorations of how to improve the performance of ANNs in classification of souvenir designs. Seven models yielded different performances. The final performances were measured by the total score.

After having all experiments done, each model was ranked with its total score again. The champion was SelBal46 which was constructed with tansig in hidden layer, logsig in output layer, balanced training with 46 successful designs and 46 unsuccessful designs in the training set, three parts of observations; the training set, validation set; and testing set with equal observations in the validation and testing set. It should be noted here that in the validation set, there were not balanced between successful and unsuccessful designs.

The score of all models as well as the ranking was presented in the table below. Some important conclusions could be made which mentioned after the table.

Table 38 Total score with descending sort

Rank	Models	Transformation function	Parts	Balance Training	Training: Validation: Testing	Total score
1	**SelBal46**	**tansig,logsig**	**3**	**Balanced46:46**	**92:50:50**	**27.84**
2	Bal46	tansig,logsig	3	Balanced46:46	92:50:50	24.70
3	Balance55	tansig,logsig	2	Balancd 55:55	110:50:50	15.09
4	FirstModel101	tansig,tansig	2	Unbalanced	235:00:50	-3.40
5	Stage4	tansig,logsig	3	Unbalanced	185:50:50	-3.87
6	3parts	tansig,tansig	3	Unbalanced	185:50:50	-6.99
7	FirstModel99	tansig,tansig	2	Unbalanced	235:00:50	-8.66

Source: Calculation using Excel

5.1 Logsig was better than tansig in the transformation function in the output layer. Models ranking 1 – 3 used the logsig.

5.2 Balance training was better than unbalanced training. Models ranking 1 – 3 also used balance training.

5.3 Division of observations into three parts; training set, validation set, and testing set showed better performance over two parts which included only training set and testing set. Models ranking 1 – 2 were constructed with three parts.

5.4 Selective learning (Komsan 2006c), including only significant variables from Logit model (see appendix) into ANNs, yielded better performance over other models with all input variables.

5.5 Selective learning contributed to the reduction of the number of neurons in the hidden layer. Only 10 neurons were needed for the better performance over many more neurons in other models.

6. Acknowledgments

The authors are grateful to Assoc. Prof. Krisorn Jittorntrum for the Neural Networks and the intellectuality he taught and brought to both authors.

7. References

Komsan Suriya. 2006a. **Tools for Prediction of Souvenir Sale in International Tourist Market using Neural Networks**. Mimeo cited in Komsan Suriya and Siriporn Srichoochart (2007). The Analysis and Development of Tourism Products and Services for SMEs and Community Enterprises in Lanna Provinces. Chiang Mai: Social Research Institute, Chiang Mai University. (in Thai)

Komsan Suriya. 2006b. Forecasting Crude Oil Price using Neural Networks. **Chiang Mai University Journal 5**, 3 (September – December) pp. 377 – 385.

Komsan Suriya. 2006c. **Airline Market Segments after Low Cost Airlines in Thailand: Passenger Classification using Neural Networks and Logit Model with Selective Learning.** Proceeding presented in the 12th Asia Pacific Tourism Association and 4th Asia Pacific CHRIE, June 26th -29th , 2006, Hualien, Taiwan.

8. Appendix (Estimation result from Logit model)

Table A-1 Estimation result from Logit Model with 18 input variables
(input values are in -1 and 1 while output variable is 0 and 1)

Dependent Variable: GOODSALE
Method: ML - Binary Logit (Newton-Raphson)
Date: 01/21/07 Time: 11:33
Sample(adjusted): 1 285
Included observations: 285 after adjusting endpoints
Convergence achieved after 6 iterations
Covariance matrix computed using second derivatives

Variable	Coefficient	Std. Error	z-Statistic	Prob.
HANDMADE	-0.130889	0.279695	-0.467970	0.6398
WOOD	-0.010733	0.235618	-0.045553	0.9637
CLOTH	-0.264637	0.261334	-1.012637	0.3112
SILVER	0.350067	0.358220	0.977240	0.3285
NATRMATR	0.141499	0.233747	0.605350	0.5449
NATURE	0.269072	0.219869	1.223781	0.2210
LANNA	0.633410	0.482300	1.313311	0.1891
THAI	0.294403	0.548402	0.536837	0.5914
ELEPH	0.659149	0.361030	1.825743	0.0679
SMALL	0.478204	0.253946	1.883094	0.0597
LIGHT	-0.201218	0.520824	-0.386346	0.6992
TOLERENC	1.144355	0.576033	1.986612	0.0470
ROUND	-0.131593	0.357420	-0.368174	0.7127
FOLD	0.258594	0.237943	1.086790	0.2771
PACK	-1.356422	0.681138	-1.991405	0.0464
USE	0.994951	0.705971	1.409336	0.1587
BODY	0.056339	0.211711	0.266115	0.7902
COLLECT	1.548287	0.678964	2.280367	0.0226
C	-0.506664	1.031507	-0.491187	0.6233

Mean dependent var	0.224561	S.D. dependent var	0.418027
S.E. of regression	0.406350	Akaike info criterion	1.077820
Sum squared resid	43.92207	Schwarz criterion	1.321319
Log likelihood	-134.5893	Hannan-Quinn criter.	1.175433
Restr. log likelihood	-151.7969	Avg. log likelihood	-0.472243
LR statistic (18 df)	34.41521	McFadden R-squared	0.113359
Probability(LR stat)	0.011188		

Obs with Dep=0	221	Total obs	285
Obs with Dep=1	64		

Source: Calculation using Eviews

Research Accreditability Report

Type: Research

Title: **The Exploration of Classification using Neural Networks: Case Study of Souvenir Designs**

No.	Name of Researchers	Address	Credit (%)
1	**Assistant Prof. Komsan Suriya**	Faculty of Economics Chiang Mai University 239 Huaykaew Rd.,T. Suthep, A. Muang, Chiang Mai 50200, Thailand Phone: +66(0)53-94-2238 Fax: +66(0)53-89-2649 Email: skomsan@econ.cmu.ac.th	**50**
2	**Mr. Sumate Prukrudee**	Social Research Institute Chiang Mai University 239 Huaykaew Rd.,T. Suthep, A. Muang, Chiang Mai 50200, Thailand Phone: +66(0)53-94-2583 Fax: +66(0)53-89-2649	**50**

……………………..…………………… ………………..………………………..

(Mr. Sumate Prukrudee) **(Assistant Prof. Komsan Suriya)**

Chiang Mai, Thailand
November 19th, 2006